TRANSLATION NO. 2239

DATE: Oct 1967

DDC AVAILABILITY NOTICE

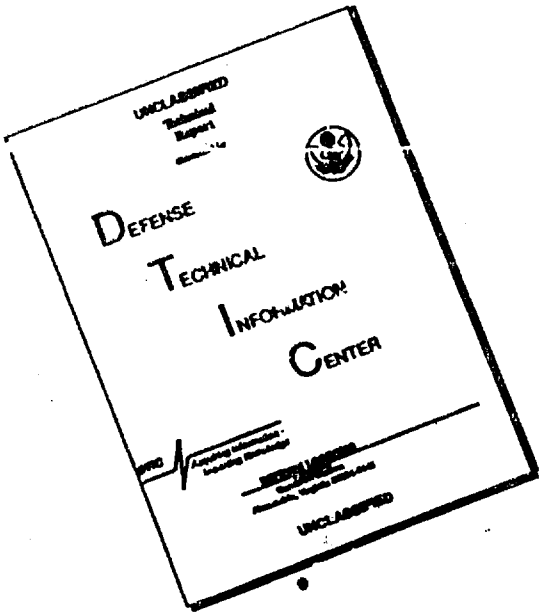
Qualified requestors may obtain copies of this document from DDC.

This publication has been translated from the open literature and is available to the general public. Non-DOD agencies may purchase this publication from the Clearinghouse for Federal Scientific and Technical Information, U. S. Department of Commerce, Springfield, Va.

DEPARTMENT OF THE ARMY
Fort Detrick
Frederick, Maryland

B

DISCLAIMER NOTICE



THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

STATISTICAL METHODS OF ANALYZING THE SEASONAL CHARACTER OF DISEASES WITH A DROPLET MODE OF INFECTION

(Following is the translation of an article by B. L. Shernasskiy and N. V. Likhnevich, Central Scientific-Research Institute of Epidemiology, USSR Ministry of Public Health, published in the Russian-language periodical Zhurnal Mikrobiologii, Epidemiologii i Immunobiologii (Journal of Microbiology, Epidemiology and Immunobiology) No 11, 1966, pages 115--122. It was submitted on 21 April 1966. Translation performed by Sp/7 Charles T. Ostertag Jr.)

The problem concerning methods for studying the seasonal nature of infectious diseases continues to remain debatable. Existing definitions of the concept of seasonal nature are contradictory in many respects.

In the literature it is possible to encounter a very diverse application of the term "seasonal fluctuation." Some authors consider it as an increase in the level of morbidity in a specific season of the year, others recognize it as equivalent to the concept of seasonal rise, and finally, the intra-annual distribution of cases by months is also sometimes called seasonal fluctuation.

For the subsequent treatment of rational methods for the quantitative characterization of the intra-annual distribution of morbidity it is necessary to differentiate the concepts of "seasonal fluctuation", "seasonal rise" and "level of morbidity, conditioned by seasonal fluctuation."

Apparently it is most correct to view seasonal fluctuation as a property which is objectively inherent to the epidemic process in the majority of infectious diseases, and which is expressed by a regularly repeating increase of morbidity during specific months or seasons of the year, caused by the influence of seasonal factors. Another reflection of seasonal fluctuation is the period of the seasonal decline, during the beginning of which morbidity usually continues to remain at a significant level, but reaching a minimum afterwards. Thus the result of seasonal fluctuation is the formation of a high level of morbidity in specific periods of the year, made up from cases belonging to the period of the rise and to the period of the decline. In its turn, it is mainly the level of morbidity in these periods which determines the height of the annual level of incidence.

Seasonal fluctuation is manifested under the influence of the so-

seasonal factors, that is, factors which have a regular periodic influence on the epidemic process during specific months or seasons of the year.

In the present report we will examine the methodical aspects of analyzing the intra-annual distribution of incidence.

When studying the intra-annual distribution of incidence by months, an analysis is made of the absolute or average daily number of cases by month, the relative significance of the number of cases for each month in the annual total of cases (relative numbers of distribution), the ratios of the average daily indices for each month to the January index or to the minimum in that year, or to the average daily index for the year (graphic indices).

The use of these methods is fully justified when studying the nature of the monthly distribution of incidence for a given year. Along with this, an analysis of each of these methods leads to a conclusion concerning their limited suitability for epidemiological practice, which has the goal of analyzing the intra-annual distribution of incidence by months not only for a given year and a given territory, but also in a comparison for various territories and for various years.

The percentage ratios of the monthly number of cases to their yearly total (specific proportion), just as any extensive indices, demonstrate only "how the phenomenon under study is apportioned into its component parts" (Merkov and Chaklin, 1962), that is, they characterize the structure of the intra-annual distribution of incidence, without answering the question concerning the frequency (intensity) of incidence in individual months. In addition to this, changes in the relative significance of the number of cases in individual months in different years may be caused not only by a true change in the number of cases in these months, but by a reverse change in the number of cases in the course of one or several other months.

Consequently, a comparison of the distribution of incidence during specific periods of the year in the dynamics over a number of years on the basis of the stated method of calculating the intra-annual distribution of incidence may lead to a false conclusion. Based on this consideration, the analysis of the intra-annual distribution of incidence on the basis of graphic indices turns out to be little suitable.

When analyzing the intra-annual distribution of incidence on the basis of the absolute or the average daily number of cases by month, consideration is not given to differences in the strength of the population in various territories, and even for the same territory in various year. Therefore, differences in the intensity of seasonal rises which are frequently detected on the basis of such indices may in actu-

ality be connected only with differences in the numerical strength of the population.

It follows from what has been said that for the purpose of analyzing the intra-annual distribution of incidence in a comparison for a number of years and based on various territories, it is more expedient to calculate the monthly incidence rate in coefficients of intensity, that is, in indices per 10,000 or 100,000 of population.

It should also be taken into consideration that the rates of seasonal rises in morbidity on the same territory in various years as well as on various territories are determined to a significant degree by the age composition of the population. In the various age groups of the population there may be different indices of morbidity, sometimes in the absence of actual differences in the number of cases. Therefore, for a comparative evaluation of the intra-annual distribution of incidence for various territories or for the same territory in various years it is necessary to analyze this phenomenon, taking into consideration the differences in the age structure of the population for the years or territories being compared. For this it is expedient to calculate the intra-annual distribution of incidence in indices per 10,000 or 100,000 of population of the specific ages which are affected the most by that infection. For droplet ("children's") infections an insignificant probability of error would be found when using as this contingent the population in ages up to 15 years.

We will dwell in more detail on the method of calculating the intra-annual distribution of incidence on the basis of the ratio of the absolute number of cases for each month to the average monthly number of cases for the year (or the ratio of the average daily number of cases for each month to the average daily number of cases for the year). The fact is that an inaccurate concept of "seasonal rise" of incidence is often connected with this method of calculating the intra-annual distribution of morbidity. Authors, using this method, refer to the period of seasonal rise as those months when incidence exceeds the average yearly level. We will examine this position in several examples.

Figure 1 depicts the intra-annual distribution of whooping cough incidence in the Latvian SSR for 1958 and 1962. In 1958, incidence exceeded the average annual level in July--August and November. With the stated approach to the determination of seasonal nature, it is expressly these months which should be considered as the period of the seasonal rise in incidence. However, in this case the period of the seasonal rise did not include May and June, when a clearly expressed increase in incidence was observed, though its level still did not exceed the average yearly index. At the same time the period of the seasonal rise includes August, when incidence, as a result of the preceding rise, was found at a level, exceeding the average yearly level, but displayed a clear tendency toward lowering.

Still more graphically revealing the imperfection of such an approach to the determination of seasonal rise is an analysis of the intra-annual distribution of incidence in 1962. In this year incidence exceeded the average yearly level in March and October--[illegible] during the period of the usual summer seasonal rise in [illegible] (in this case June--July) the level of incidence did not exceed the average yearly index, and consequently the summer rise was not [illegible].

Other deficiencies in analyzing the monthly distribution of incidence based on the relationship to the average yearly level are revealed in an example of the intra-annual distribution of measles incidence in the Azerbaydzhan SSR in 1953--1955.

Figure 2 distinctly shows the fall-winter seasonal nature inherent to measles, characterized in this case by the fact that the seasonal rise begins at the end of one year and ends in the beginning of another. At the same time, an analysis of the monthly distribution of incidence in respect to the average yearly level, limiting the study of the dynamics of morbidity within the framework of one year, permits only segments of the seasonal curves to be taken into consideration. In one case this is the termination in the rise in morbidity which had begun in the previous year, and in the other case - the beginning of a rise ending in the following year.

Thus, the method of studying the intra-annual distribution of incidence on the basis of the relationships of the monthly levels to the average monthly level for the year turns out to be little suitable for analyzing the seasonal rise, since the seasonal rise begins before incidence exceeds the average monthly level and is completed before incidence becomes lower than the average monthly level. Consequently, this method does not allow consideration to be given to the beginning of the rise, at the same time making it necessary to relate to the period of rise those months when the rise in incidence is replaced by a decline. Besides this, the use of the stated method makes it necessary to assign to the period of seasonal rise those months when incidence exceeded the average yearly level, not due to an increase in the influence of seasonal factors, but on the strength of accidental cases, while the truly seasonal rise may remain unconsidered due to the fact that incidence during this period might not have exceeded the average yearly level.

The question of what period should be considered as the period of the seasonal rise can be answered most correctly, as it seems to us, if we accept the point of view of those authors (Guslits, 1959; Rogozin and Shura-Bura, 1965) who calculate the period of the seasonal rise from the month when incidence begins to increase in comparison with the minimum level up to the month with the maximum level (after which a drop in incidence begins, lasting to the following minimum level).

However, it must be stipulated that the month in which incidence begins to increase in comparison with the minimum level cannot always be considered as the onset of the seasonal rise. It is necessary to be convinced first of all that the prevalence of incidence in this month in comparison with the previous one is not an accidental occurrence, but is actually conditioned by the influence of some additional factors. An explanation of this circumstance in respect to the so-called standard seasonal curve, obtained on the basis of [illegible] data from several years, can be furthered by the calculation of the statistical reliability (or as they say, significance) of the difference in the morbidity indices during the months of interest. Correspondingly with this onset of the period of seasonal rise it is necessary to take into consideration the month when incidence essentially exceeds the level of incidence during the previous months.

Evaluation of the significance of the various morbidity indices in this case may be carried out either by means of comparing the difference of the indices with the average error of difference, or by means of determining the reliable interval on the basis of binomial distribution. The proposed methods for clarifying the true onset of the seasonal rise make it possible, when evaluating the dynamics and level of incidence during the period of the rise, to exclude from this period the months when the increase of incidence turns out to be accidental, not caused by the influence of seasonal factors.

However, statistical methods should be used in conjunction with methods of epidemiological observation and reporting, they should concretely define the latter and make them more precise. Therefore, as an example, an increase in incidence, established as a result of epidemiological observation, and known to be conditioned by the influence of seasonal factors, should be accepted by the epidemiologist. Statistical methods should be called for only to give a quantitative evaluation of the significance of this growth.

Both quantitative and qualitative indices are used for characterizing the seasonal nature of infectious diseases.

Qualitative indices may include reports of the months for the onset and termination of the seasonal rise, and also descriptions of the nature of the morbidity curve during the period of the rise (for example, the seasonal rise began in such a month, terminated in such a month, incidence during the period of the seasonal rise increased uniformly or in individual months its dynamics acquired a spasmodic nature).

Since the seasonal curve reflects the process of a rise in incidence from month to month, the requirement arises for the quantitative characterizing of the dynamics of morbidity during the period of the seasonal rise.

For evaluating the seasonal rise, use is sometimes made of Herring's so-called index of the seasonal nature, which represents the relationship of the number of cases in the months with the maximum and minimum level

However, this index is little suitable for [illegible] the seasonal curve, since it does not reflect either the nature [illegible] of morbidity for the entire period of the rise or the duration of this period.

Also little suitable for evaluating the seasonal curve for the same reason is the calculation of the average geometric rates for the increase in incidence according to the formula:

since in this formula also consideration is not given to the nature of the dynamics of morbidity for the duration of all the months of the seasonal rise.

The dynamics of morbidity for the course of two contiguous months is clearly characterized by the index of the relative rapidity in the tempo of increase in incidence in each subsequent month in comparison with each preceding one. It is determined based on the difference between the subsequent and preceding level of an order (absolute increase), relative to the preceding level:

Such a methodical approach to evaluating the increase in incidence in individual months makes it possible to find the path to the generalized characteristics in the dynamics of incidence throughout the entire period of the seasonal rise: Most fully characterizing the dynamics of development of morbidity during the period of the seasonal rise is the index of average increase in morbidity for the certain period, calculated, taking into consideration the values of each of the members of the succession, or, in other words, the index of the average monthly tempo for the dynamics of incidence for the period of the seasonal rise. The index of the average increase in incidence during the period of the seasonal rise represents the generalized nature for the tempo of change in the given phenomenon for the period of time being studied. This index can be obtained by the formula for the weighted mean geometric coefficient of growth:

In this way this formula may be written in the following form:

where - is the index of the average monthly tempo in the dynamics of incidence during the period of the seasonal rise, - incidence in the month, preceding the seasonal rise (minimum), , ... - incidence in each of the months during the period of the seasonal rise, - incidence in the last month of the seasonal rise (maximum), - duration of the seasonal rise (from to).

For the convenience of calculation the operation of evolution if replaced by the use of logarithms:

In the end calculation we obtain the average tempos of increase, expressed in fractions of a unit. For expressing this number in percentages it should be multiplied by 100.

As an example, in Figure 3 we will present the indices for the average monthly tempos of change in scarlet fever incidence in the Moldavian SSR in 1950 and 1960. In 1950 the minimum index of morbidity was in April, the seasonal rise in incidence began in May, continued for 6 months, and reached a maximum in October. In 1960 the minimum index was in July, the rise began in August, continued for 3 months and reached a maximum level in October. In both cases the ratio of the maximum and minimum incidence turned out to be similar: In 1950 - 5.9, and in 1960 - 5.3 Calculation of the average geometric tempos of increase in incidence according to the formula

made it possible to calculate the differences in the duration of the seasonal rise in the years being analyzed. In connection with the fact

that in 1950 the rise lasted 6 months and in 1960 - 3 months, the average tempos of increase in incidence during the periods of seasonal rises for these years turned out to equal correspondingly 0.347 and 0.744, or 34.7 and 74.4%, a month, that is, in 1960 the incidence during the period of the seasonal rise increased from month to month more intensively than in 1950.

The development of incidence during the periods of the seasonal rises for these years can be characterized still more accurately by a calculation of the index for the average tempo of increase in incidence, taking into consideration the value of each of the members of the series: In connection with the slow increase in incidence in the beginning of the seasonal rise in 1950, the index for the average increase in incidence equaled 0.173, or 17.3%, while in 1960 it equaled 0.738, or 73.8%.

Thus, the index for the average tempo of increase makes it possible to characterize the dynamics of the seasonal rise in incidence, taking into consideration the rate of development of this process and its duration.

As a criterium for a quantitative evaluation of the level of incidence during the period of the seasonal rise, and also for the drop in incidence, it is possible to use the intensive indices of incidence during these periods (indices of intensity for the seasonal rise or drop), calculated per 10,000 or 100,000 population. The index of intensity for the seasonal rise is calculated bymultiplying the total of incidences during the months of the seasonal rise by 10,000 or 100,000 and dividing the product by the average strength of the population over the period of the rise.

It should be kept in mind, however, that the level of incidence during the period of the seasonal rise interests the epidemiologist mainly from the point of view of its influence on the overall annual level of incidence. Therefore, besides obtaining the direct characteristics for the level of incidence during the period of the seasonal rise, it is necessary to evaluate this level in relation to the annual level of incidence. For this purpose the index of intensity for the seasonal rise is divided by the index of incidence for the year. It is understandable that the closer this ratio comes to a unit, the greater the significance of the seasonal rise in the forming of the incidence level for the year.

Sometimes the calculation of the intensity index for the seasonal rise and the intensity index for the year is somewhat complex, in connection with the fact that the seasonal rise in incidence begins in one year and ends in the next year, while during the given year the termination of the seasonal rise which had begun in the previous year takes place. Such a phenomenon is often observed for measles. In similar

cases the index of intensity must be calculated based on the sum of incidences for those months of that year during which the termination of a rise which had began in the previous year takes place, and for the months for the onset of the seasonal rise for that year. After this the index of intensity for the seasonal rise, calculated in this manner, is divided by the index of incidence for the year.

Conclusions

1. For analyzing the intra-annual distribution of incidence in comparison with a number of years and for various territories, it is expedient to calculate the monthly incidence in indices per 10,000 or 100,000 of population from specific age contingents.

2. A reliable determination of the onset of the period of the seasonal rise in incidence can be obtained by a statistical analysis of the intra-annual distribution of incidence on the basis of evaluating the significance in the difference of indices.

3. For a quantitative evaluation of the development of incidence during the period of the seasonal rise it is possible to use the indices for the average tempo of dynamics of incidence, calculated by the formula of the weighted mean geometric coefficient of increase in incidence, taking into consideration the indices of incidence in each of the months of this period. This same method may be used for the quantitative characteristics of the dynamics of incidence during the period of the seasonal decline.

4. As a criterion for the quantitative characteristics of incidence levels during the period of the seasonal rise, and also decline, in incidence, it is possible to use the intensive indices of incidence during these periods.

Literature

1. Guslits, S. V., In the book: Theoretical Problems of Epidemiology, Kiev, 1959, p 112.

2. Merkov, A. M., Chaklin, A. V., Statistical Study of Malignant Growths, Moscow, 1962.

3. Rogozin, I. I., Shura, Bura, B. I., In the book: Multi-volume Textbook on the Microbiology, Clinical Aspects and Epidemiology of Infectious Diseases, Moscow-Leningrad, 1965, vol 5, p 151.

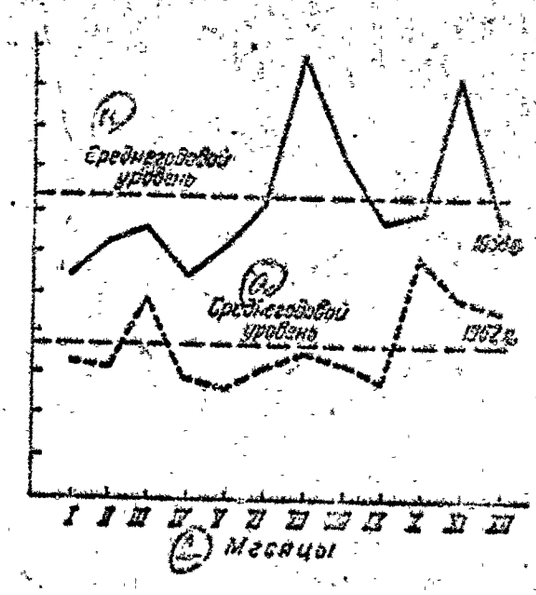


Figure 1. Monthly distribution of whooping cough incidence in the Latvian SSR in 1958 and 1962.

a - average yearly level; b - months.

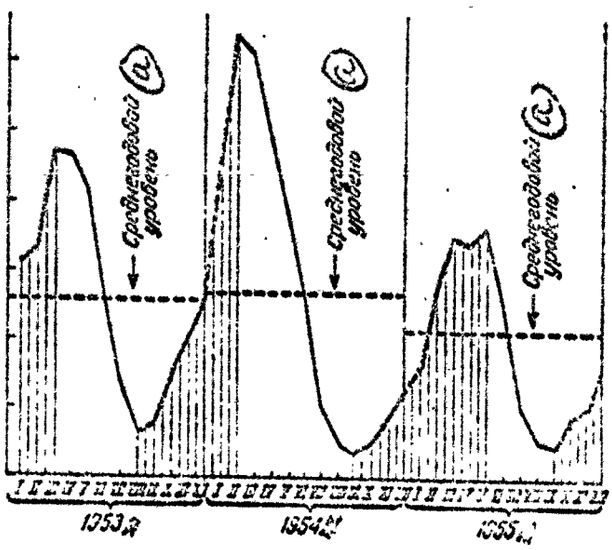


Figure 2. Monthly distribution of measles incidence in the Azerbaydshan SSR in 1953--1955.

a - average yearly level.

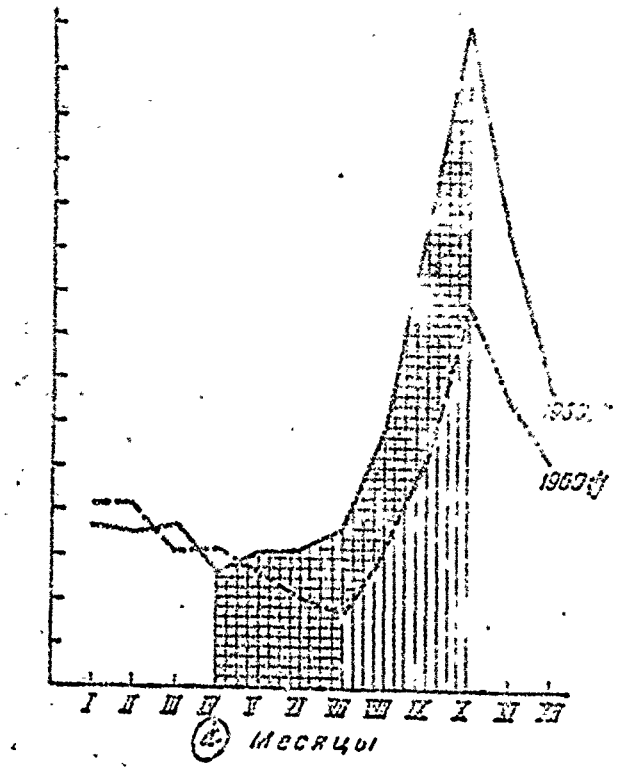


Figure 3. Average monthly tempos of change in scarlet fever incidence in the Moldavian SSR in 1950 and 1960.

a - months.